Educ.

EDUCATIONAL RESEARCH BULLETIN: ERB

DUKE UNIVERSITY LIBRARY
NOV 26 1958
DURHAM, N. C.

NOVEMBER 12, 1958

COLLEGE OF EDUCATION
THE OHIO STATE UNIVERSITY

Vol. XXXVII, No. 8

EDUCATIONAL RESEARCH BULLETIN

(Entered in the United States Patent Office)

R. H. ECKELBERRY.....*Editor*
SHIRLEY LIPSON.....*Assistant Editor*

PUBLISHED BY THE BUREAU OF EDUCATIONAL RESEARCH AND SERVICE, OHIO STATE UNIVERSITY, FOR THE SCHOOLS OF OHIO. EDITORIAL OFFICE, ROOM 173; BUSINESS OFFICE, ROOM 195, ARPS HALL, COLUMBUS 10, OHIO

ANNUAL SUBSCRIPTION ONE DOLLAR ($1) TO PERSONS LIVING OUTSIDE OHIO

CONTENTS

The Student Personnel Program in the College of Education 197
COLLINS W. BURNETT, *Ohio State University*

The Development of Pupil-Growth Criteria in Studies of Teacher Effectiveness 205
HAROLD E. MITZEL AND CECILY F. GROSS, *College of the City of New York*

Acceptance and Performance among Gifted Elementary-School Children 216
META F. WILLIAMS, *University of Bridgeport*

Editorial Comment 221

Readings 223
The Child and His Elementary School World.

THE EDUCATIONAL RESEARCH BULLETIN is published monthly except during June, July, and August. Entered as Second-Class Matter, February 9, 1922, at the Post Office at Columbus, Ohio, under the Act of August 24, 1912. Accepted for mailing at specified rates of postage provided for in Section 1103, Act of October 3, 1917, authorized February 21, 1922.

Educational Research Bulletin

NOVEMBER 12, 1958
Vol. XXXVII, No. 8

The Student Personnel Program in the College of Education

By COLLINS W. BURNETT

THE student personnel program in the College of Education at Ohio State University has undergone many modifications since 1928 when George W. Rightmire, then president of the University, set up the position of Junior Dean in each of the five undergraduate colleges. The Junior Dean's main function was to serve as a student personnel administrator with especial concern for the problems of the freshman and sophomore students. Now, thirty years later, even the title of Junior Dean has been discarded for more impressive ones such as Co-ordinator of Student Personnel and Assistant Dean. To some people, the "student personnel program" suggests a point of view or attitude which establishes a friendly, permissive climate for students; to others, the phrase means personnel who work closely with students in a non-textbook relationship; and to still others, the term means techniques such as counseling, testing, and the use of occupational information. In reality, all three of these concepts—philosophy, personnel, and techniques—are valid ones to use in describing the student personnel program in the College of Education.

Any philosophical considerations for the College of Education have to be determined within the larger context of a land-grant university. Historically, there has been little emphasis on selective admissions procedures. Generally speaking, any graduate of a four-year high school in Ohio has been eligible for admission to the University and enrollment in one of the five undergraduate colleges. More than 250 programs of study are available to students who for the most part are residents of Franklin County and northeastern Ohio, areas which have become highly industrialized during the last twenty years.

The distinctive function of the College of Education is to pre-

pare teachers. Although some of the students are enrolled in three other areas—professional art, professional music, and occupational therapy—the main stream of activity is teacher preparation; hence the student personnel program in the College of Education has an underlying basis which makes it different from other student personnel programs: the welfare of the individual student becomes secondary in importance to what is best for children in the public schools. In all of the counseling relationships, the first consideration is not that the client wants to become a teacher but rather that he needs help in determining whether he meets the criteria (scholarship, speech and hearing, proficiency in written English, professional rating, and health—both physical and emotional) for becoming a good teacher.

A second distinct characteristic of the philosophy of the student personnel program is the development of a close, personal relationship between students and faculty. This is indeed important since many of these young people will become teachers and will tend to adopt some of the same personnel attitudes toward their students which they encountered when they were students.

Still another philosophical consideration is the importance of providing a psychological climate in which there is a balance between sufficient student personnel services and adequate opportunity for students to develop a sense of maturity and self-direction. Instead of viewing education as a dynamic process with unlimited opportunity for creativity and intellectual curiosity, many students tend to settle into a passive acceptance, wanting to be told when to study, which activities to join, and which courses to schedule. In developing a student personnel program, it is important not only to determine student needs, but also to decide which needs can be met by faculty and student personnel efforts and which ones should be met by students themselves.

WHO are the people who are responsible for the student personnel program in the College of Education? In a general way, the student personnel program is everyone's responsibility: the Dean and his administrative staff, the faculty, and the staff in the College Office and departmental offices where many students form their first impressions of the College. At the University level, which will be discussed later, there are also key people who contribute to the College's program of services for students.

By far the largest group of people who are interested in student welfare are the faculty who have the advantage of know-

ing about absence, poor scholarship, and behavioral problems. The College librarian, for example, on numerous occasions has performed a guidance function with students by helping them locate and develop resources for a term paper, or find a congenial rooming house; by arranging for a loan of money which is provided by a professional group; by providing a part-time job in the library; and by referring a student who is emotionally upset to the student personnel staff. Most of the faculty either because of professional training or sensitivity to student concerns have a strong feeling of responsibility in working closely with a student who by his behavior indicates some need. If the problem becomes serious enough to suggest the need for specialized professional attention, the student is referred to one of the counselors on the student personnel staff.

Two of the key groups of the faculty who regularly render outstanding student personnel service are the advisers and the co-ordinating advisers. In each department of the College there are faculty members designated as advisers who are responsible for working closely with students. They evaluate students in terms of probable teaching success, help plan course schedules, recommend approval of student petitions, recommend students for teaching positions, or refer students for professional help. Many times due to the adviser's alertness a student problem such as absence from class is spotted early enough so that the adviser and the student can work on the problem before it becomes critical.

The co-ordinating advisers, as the name suggests, are a group of faculty members who co-ordinate advising, scheduling, and other personnel matters in their respective areas. This group, which was organized about 1938, meets twice a month to discuss problems, learn about new developments in the College and the University, and suggest changes in personnel procedures. In the case of the School of Fine and Applied Arts and the School of Music, for example, where many personnel procedures are autonomous, the two co-ordinating advisers perform a heavy personnel function in addition to some teaching responsibility. Not only is each of these persons a leader in the development of the student personnel program in his own area, but he also assists in the orientation of new faculty members and co-ordinates information and problems which arise in his area with the student personnel staff of the College.

From one standpoint, the Co-ordinator of Student Personnel

and his staff function as the CIC (Combat Information Center—a Naval term designating the nerve center aboard a combat ship where all communications are controlled and incoming data about enemy planes and ships are processed). The College Office serves as the control point for all of the important data concerning closed courses, new sections of courses, absence reports, grade sheets, petitions, dismissals, reinstatements, requests for point hour for honorary organizations and Selective Service, graduation requirements, permanent record cards, withdrawals, adding and dropping courses, and results of the mathematics and English placement tests. Sometimes in the press of rush periods, which seem to occur with frightening regularity, with the constant ringing of telephones, the swarms of impatient students, and the frantic requests of faculty, the CIC seems more than a figure of speech.

From another standpoint, the student personnel staff functions as a closely knit team in co-ordinating, administering, and planning the personnel program. Service to students is paramount, but service to faculty who are working with various kinds of student problems is also important. Perhaps an hour or two spent in orienting a new faculty adviser to the mysteries of the College *Bulletin* or the vagaries of campus rules and regulations may bring about a more sympathetic understanding of student behavior. Important relationships also have to be established and maintained with personnel administrators in other college offices as well as with those who are in charge of key student personnel functions at the University level.

Although there are many persons at the University level whose work relates closely to the College student personnel program, only a few will be mentioned to establish the point that both groups of personnel and the two programs—College and University—are interdependent. The common goal is to further the progress and development of the student.

Numerous student personnel relations have been developed with the Executive Dean, Student Relations, who is a member of the President's Cabinet, and his staff which includes the Dean of Men, Dean of Women, Director of the University Counseling and Testing Center, and directors of the University Health Service, Scholarships Office, and Student Financial Aids Office. A staff rather than a line relationship exists between the College student personnel people and these members of the University student per-

sonnel area. The Co-ordinator of Student Personnel has a line relationship to the Dean and the Associate Dean of the College of Education. Similar staff relationships exist between the College student personnel area and the Executive Dean, Special Functions, who is also a member of the President's Cabinet, and whose staff includes the Registrar and the University Examiner.

THE story goes that one day when William Lowe Bryan, the late president of Indiana University, was walking across the Bloomington campus, an old graduate stopped him to ask, "President Bryan, why is it that each year this university graduates so many fools?" The president replied, "Because, each year we admit so many fools." At Ohio State also the problem of graduating only qualified students arises because very little has been done in the way of selective admissions procedures (the School of Music does have a well-developed admissions program which is permitted by legislative enactment). The fact that natural selection is operating is demonstrated by research studies which show that approximately 40 per cent of those students who enter the College as first-quarter Freshmen are not around to be counted by the beginning of the fourth quarter. Moreover, about 65 per cent of the graduating Seniors did not start in the College as first-quarter Freshmen.

The College at one time had an admissions program, but the trend in recent years has been to develop a retention program, a set of criteria to help the student evaluate his progress. This program of progressive retention which operates in all the teacher-preparation areas is implemented, first, by Education 408, formerly Education Survey 407,[1] which is intended as an orientation and introduction to the study of education. Second, Provisional Acceptance for Teaching, which must be met before student teaching can be scheduled, consists of satisfactory clearance on health, speech and hearing, professional rating, a minimum cumulative point-hour ratio of 2.0, and an average of *C* for the three courses in the freshman English sequence. Third, student field experience permits the evaluation of students in September Field Experience, leadership experience in selected community agencies, and student teaching which is scheduled in the senior year.

[1] Burnett, Collins W., and Seeman, Alice Z. "Guidance of College Freshmen," *Journal of Teacher Education*, III (March, 1952), pp. 19–22.

Any discussion of the scheduling and advisory programs points up the complexity and diversity of the student personnel program; scarcely any part of the program can be discussed without several qualifications and parenthetical statements. The Schools of Music and Fine and Applied Arts and the Department of Occupational Therapy, however, have developed techniques for scheduling and advising which, while they are related to those in the other areas of the College, are somewhat independent.

Except for one very large area in teacher education, helping the students plan their schedules of course work each quarter is a function of the advisers in various areas. Partly as an effort to provide scheduling assistance to students who are in elementary education, a plan for a Consultant's Day was developed which provides for faculty advisers to be available on an announced date so that students may go to a certain designated conference location or to the advisers' offices to get help with scheduling problems.

In a general way, the advisory program for most Freshmen who are in teacher education starts with Education 408, because when students schedule this course in the Autumn Quarter the instructor also serves as the adviser. Members of the student personnel staff serve as advisers to those Freshmen who cannot schedule the course until a later quarter. Following the freshman year, the student is assigned to a permanent adviser in his major area at the time of the "declaration interview" which must be initiated by the student between the second and fourth quarters. Continuity in the advisory program, therefore, is established for the student from the first quarter of enrollment through graduation.

The declaration interview, which is a recent development, is an intriguing "package" plan in which a one-hour interview with the student at the time he declares his major area provides the following: assignment to an adviser, completion of an area sheet which shows the courses completed and those to be completed before graduation, a review and planning of student field experience, and a review and explanation of the retention program, Provisional Acceptance for Teaching. This latter program permits the student the freedom of at least nine quarters in which to demonstrate his competence—if he started in the College as a first-quarter Freshman. Regardless of when the student enrolled in the College, he must meet all five requirements of the Provisional Acceptance for Teaching before he registers for student teaching, which cannot be done until the senior year.

Counseling is central in the student personnel program. Students are referred for counseling to a member of the College student personnel staff, perhaps by an instructor or adviser, or the contact may be self-initiated. Counseling problems are most frequent in such areas as low scholarship, challenging the superior student, indecision about vocational choice, emotional adjustment, and difficulty in meeting some College or University requirement for graduation. Counseling is primarily in the pattern of a series of one to three interviews in which screening, diagnosis, and clarification help the student reach his own decision. In many cases, the nature of the interview can be described as one in which information is presented or some confusion about a regulation is clarified. When the counselor decides that the problem of emotional upset is serious, he refers the student to a psychiatrist in the University Health Service or in less severe cases to a counselor in the University Counseling and Testing Center or the Mental Health Clinic. Neither at the College nor at the University level is there adequate opportunity for prolonged therapy with severely regressed behavior cases, which suggests that such a student can be helped best by withdrawing him from the University and helping him arrange for private treatment.

Although the Registrar's Office maintains the official student record from which transcripts are made, the College Office is responsible for the student personnel folder which contains the IBM permanent record card with the quarter-by-quarter listing of courses, marks, cumulative point-hour ratio, and entrance test scores; copies of letters sent to the student or to the adviser; petitions; counseling notes; letters of recommendation to prospective employers or to agencies for scholarships; changes of course; Selective Service Ratings; and absence reports.

Because of the democratic philosophy of the College, many policies and decisions which affect the student personnel program are established by committees, and one or more students are important participants in most of the major ones. The Committee of Co-ordinating Advisers, which is composed of key advisers from all areas of the College, discusses problems and suggestions pertinent to the advisory program. The Committee for the Student Personnel Program initiates procedures and plans and implements policy in all areas of student personnel. The Petitions Committee, which meets about once a week, reviews and decides on all petitions for reinstatement or requests for some

departure from established procedure. About once each quarter the Committee for Degrees with Distinction has the pleasant task of reviewing records of superior students (those with a cumulative point-hour ratio of 3.5 or above) who are graduating to decide which ones to recommend to the faculty for degrees with *cum laude* or *summa cum laude*.

Developing an active participation in College affairs on the part of the students is another aspect of the student personnel program. In addition to the numerous student honoraries in the various departments, there are clubs which represent specialized interests, such as the Association for Childhood Education in the area of elementary education.

The Education Student Council, which is partly honorary but is also an important service group representing the entire College, is one of the key student groups to improve relationships between students and faculty. In addition to nominating student members for most of the important College committees, the Council contributes to the development of leadership and responsibility throughout the year. During the 1957–58 academic year, the Council planned the orientation meeting for new Freshmen, initiated a series of fireside chats with students and faculty, planned a recognition tea for Freshmen with scholarships, and organized an all-day conference of students and faculty to evaluate all aspects of the College program.

Although scholarships, loans, and part-time jobs are aspects of student personnel work which are centralized for the most part in University-level offices, the College takes significant action in these areas also. In addition to helping select one hundred Freshmen for teacher-education scholarships, the student personnel staff recommends many other students each year for general University scholarships. Numerous loans are effected through the help of a rotary fund established by Phi Delta Kappa. Some part-time jobs are also "discovered" and made known to students whose need has become clear through counseling.

Although student teaching and placement have strong personnel connotations,[2] both areas will only be mentioned because each is administered outside the student personnel program. Its staff, however, does contribute significantly to them.

[*Continued on page 224*]

[2] Dickson, George E., and Burnett, Collins W. "A Student Personnel Approach to Student Teacher Placement," *Educational Administration and Supervision*, XLI (November, 1955), pp. 390–96.

The Development of Pupil-Growth Criteria in Studies of Teacher Effectiveness

By HAROLD E. MITZEL AND CECILY F. GROSS

NOTE.—In the October issue of the BULLETIN, the authors presented a brief review of twenty research studies dealing with the use of pupil growth as an aid in the assessment of teacher effectiveness. These studies covered the educational literature of a thirty-year period, from 1921 to 1951. In the second section of the paper, which follows, a number of fundamental questions are raised concerning the philosophy, conduct, and analysis of research employing pupil growth as a criterion of teacher effectiveness. The reader is reminded that the bibliographic sources can be found at the end of the first section of this paper in the October issue of the BULLETIN (pp. 186–87).

AN ANALYSIS of the attempts to devise pupil-growth criteria as measures of teacher proficiency reveals serious weaknesses. In some cases the deficiencies resulted from poor administrative conditions or from the lack of funds for a maximum research effort. On the other hand, many of the studies cited in the first part of this paper could have been improved by the use of more rigorous procedures; others, by the researcher's preliminary clarification of certain fundamental issues.

On the following pages, we list some of the basic questions of logic and methodology involved in the measurement of teacher competence in promoting pupil growth. The more philosophical questions deal with the conceptualization of the problem of evaluating teaching effectiveness. Other questions deal with experimental conditions or controls for research design. A third group concerns the appropriateness of different methods of analysis and the necessity for taking into account a number of sources of variation. The researcher's position on these questions, whether explicit or implicit, inevitably influences his research design. Furthermore, recognition of the limitations placed upon the interpretation of results by imperfect controls or by the assumptions on which the statistical method is based prevents the drawing of unwarranted conclusions.

For the past fifty years, most research has been based on the assumption that teaching competence is a unitary trait. Although the degrees of effectiveness among teachers have long been realized, many educators still seem to assume that the teacher who stimulates the greatest pupil growth in one basic skill will stimulate the greatest growth in other specific skills as well as in problem-solving, social adjustment, and all other desirable outcomes. Moreover, they also seem to assume that the effective teacher is equally effective with all children. Until these assumptions are supported by empirical evidence, it is perhaps more reasonable to assume that the effectiveness of teachers may vary from goal to goal and, therefore, that research be conducted on pupils' achievement of a wide variety of different goals. Ackerman (1: p. 285) states the case for a multidimensional concept of teaching competence:

> The study of teacher effectiveness must assume the possibility of different kinds of effectiveness for different kinds of teachers, pupils, programs or situations. Inherent in the idea of effectiveness is a potential variety of "effectiveness." Research can attempt to discover the various kinds of effectiveness. It cannot tell us which is the best or most desirable. This is a judgment which each individual must make according to his own educational goals and values.

The probable multidimensional nature of teaching effectiveness assumes practical importance when one must decide whether or not to combine different pupil-growth measures into a composite labeled "effectiveness." That teachers who were effective in teaching reading were not always effective in teaching arithmetic and vice versa is illustrated by Taylor's (25) study. He found a correlation of .232 between arithmetic gain and reading gain, although the reliabilities for these growth criteria were .864 and .415, respectively. If Taylor had combined these two types of effectiveness, it would have been rather like describing an individual's appearance by the sum of his height in inches and his weight in pounds. A rule of thumb might be that two or more pupil-growth measures should *not* be combined unless their intercorrelations are of the same order of magnitude as their separate reliabilities (26).

The work of McCall (19) suggests a complete adherence to the concept of unitary effectiveness, but in the operational definition of his growth variables, he does not postulate that all growths

are identical with one another. The study of Lins (17) is seriously defective on this score. Lins compared the effectiveness of high-school teachers of biology, civics, English, social science, and general science on the basis of different standardized tests given to their respective classes in these subjects. Certainly it is illogical to regard five tests of five different subjects as measures of a single attribute. Transformation of the data into standard scores in no way meets this basic objection. The study by Bimson (5) errs in the same manner by treating scores in algebra, general science, and history as equivalent.

A rational criterion of equivalence of tests is needed. In the absence of such a criterion, we can only repeat our suggestion that unless the intercorrelations are of the same order as the reliabilities, measures of growth should not be combined. To use one test for one class and a different test for another (unless they have been shown to be equivalent) is indefensible.

The studies done by Wisconsin students under Barr's direction are not firmly committed to either the assumption of unidimensionality or the assumption of multidimensionality. Rostker's study (23) is a case in point. Rostker analyzed his eight growth measures separately, in various combinations, and as a composite of all eight. Although no statistical principles are violated, there appears to be no reason for combining the growth measures since their intercorrelations were uniformly low. The most plausible conclusion is that his different tests measured different growths which may have been stimulated in part by different kinds as well as different degrees of teacher effectiveness.

The examples cited show that the researcher's answer to the question about the nature of teaching effectiveness, which at first appeared to be purely philosophical, exerts a crucial influence upon the practical conduct of his investigation.

THE social importance of any investigation of teaching effectiveness depends upon the relative rank of the growth measured in some hierarchy of educational objectives. For this reason, attempts to measure pupil growth should be preceded by thoughtful consideration of the comparative importance of various educational goals. Presumably those growths judged most important ought to be measured first, depending upon whether suitable and adequate measuring devices exist or can be constructed. McCall's study (19) is a model as regards the attention

given to defining and ordering educational objectives for the elementary schools of North Carolina. A panel of experts assisted in choosing and weighting different areas of pupil growth.

Consistent with some value system, any logical procedure for deciding which kinds of growth to measure is clearly preferable to the method of using whatever test results happen to be available in school files. The research by Betts (4) illustrates a procedure in which Stanford Achievement Test data already on file determined the type of pupil growth that was measured. This sort of planning suffers from two major weaknesses. Unless researchers collect their data personally, rather than depend upon teachers or school supervisors to give tests to their own pupils, standardization of test administration is hard to ensure. Furthermore, if the teachers conduct the initial testing of their pupils, the possibility that some of them will teach for mastery of specific test materials becomes dangerously high.

Involved in the selection of growths for measurement is the problem of choosing adequate measuring instruments. In addition to meeting technical standards, the selected tests ought not to impose "floor" or "ceiling" effects. This limitation is especially crucial in those studies which attempt to control mental ability by means of the accomplishment quotient. Tests with too low a "ceiling" will tend to show an artificial negative relationship between growth and mental ability. As a result of this distortion, bright children with higher initial scores appear to show less gain than duller children whose initial scores were low. Crabbs (8), in particular, was aware of this difficulty, although she depreciated its influence upon the growth criterion. As noted above, La Duke (16) and other Wisconsin investigators also discuss the effect of test "ceilings."

Pupil-growth research, to be worth while, must deal with major school objectives. Therefore, a great deal of attention needs to be given to the problems of defining growths consistent with a hierarchy of educational goals, and of selecting or building devices to measure their attainment.

THE problem of deciding how the objectives chosen for measurement by researchers be related to those sought for pupils by teachers involves the distinction between a teacher's success in doing what he is trying to do and his success in doing what, for purposes of the study, the researcher thinks he ought to be doing.

It may be argued that if Teacher A believes that reading ability is the most important school objective at the elementary-school level and if Teacher B believes that "learning to work together" is the prime objective, then, on the assumption that both teachers teach according to their convictions, the reading achievement of the respective classes may be fundamentally incomparable indexes of teaching effectiveness. If, on the other hand, the research worker specifically wants a measure of effectiveness in the teaching of reading, he may consider the private-goal hierarchy of individual teachers as irrelevant.

One way of reconciling the educational objectives set by investigators with those of teachers is to measure pupil achievement of goals which are restricted in scope. Almost all elementary-school teachers stress the necessity for teaching children to read and compute, but they may differ widely in their ideas of the relative importance of economic as contrasted with physical geography. Those researchers who desire that the goals of teachers match the abilities of pupils chosen for testing will probably encounter fewer difficulties in the measurement of gain in skill areas, such as reading comprehension or arithmetic fundamentals, than in subject areas, such as geography and science. Nonetheless, the advantage derived from restricting the growth area to be measured ought not to involve a compromise with the choice of the most important educational objectives.

In the studies reviewed, very little attention was given to the relationship between teachers' objectives and the kinds of measured pupil achievement. However, centralized curriculums and courses of study on a state-wide or city-wide basis, questionable though they may be to some educational philosophers, have facilitated the research of certain investigators by providing standardized objectives. The study by Brookover (6) was based on the use of two forms of a United States history examination prepared by the Indiana State Department of Instruction. The same course of study and textbook were used by all 66 teachers and their pupils. Rostker (23) made certain that his 28 teachers had common objectives for two short-term units in social studies by specifying four or five major subtopics for each unit. The teachers were free to develop methods, materials, and content.

In whatever way research workers resolve the question of possible conflict of purpose, they ought to be aware of its implications for the interpretation of their findings.

For ideal experimental control of the effects of teachers other than those under study, growth should be measured on pupils who are taught by only one teacher. Under present-day systems of school organization, the nearest approximation to this ideal can be found in the elementary grades. Although even at the elementary-school level pupils have art, music, and other special teachers, the effect of such teachers in the lower grades can more readily be ascribed to a general school effect than at the high-school level. In a departmentalized high school, it is unlikely that all social-studies teachers, for example, will refrain from teaching language skills. Nor would it be considered desirable that they do so. Still, the efforts of social-studies teachers could bias the results of a study of the effectiveness of the English teacher as measured by pupil growth in English. Therefore, it seems easier at this stage in the development of pupil-growth criteria to impose gross experimental control by concentrating on the study of elementary-school teachers and pupils. The necessary replication at the secondary level will undoubtedly need to reckon with this source of error.

In the studies of Brookover (6), Lins (17), Jones (15), and Bimson (5), all based on high-school pupils taught by more than one teacher, no attempt was made to control the growth criteria for possible influence by teachers of other subjects.

There are two chief ways of controlling possible differences in pupil growth from one grade to another. One method requires an adequate sample of teachers and pupils from each of a number of grades. This procedure represents a project of great magnitude and complexity. An alternative method has the advantage of keeping the study within manageable proportions. A single growth study can be limited to the measurement of pupils in a single grade or a small grade range. Future replicative studies, however, would be required to answer the question whether teacher behaviors related to effectiveness are the same at all grade levels.

Restricting the grade range to be covered in a single study offers other advantages related to the problem of selecting appropriate educational objectives for measurement. The use of sample teachers from a single elementary grade is more likely to ensure that the curricular aims of all the teachers under study are as nearly alike as possible. Furthermore, by drawing classes from a

single grade, the researcher reduces the likelihood that variations in the hierarchy of importance of educational goals may be encountered. Arithmetic is an example of a skill subject for which the objectives of instruction change from grade to grade. Comparing growth in fourth-grade arithmetic with growth in eighth-grade arithmetic may yield differences primarily because of the unlike emphasis placed on arithmetic instruction in the two grades.

Investigations by McCall (19) and Stephens and Lichtenstein (24) are noteworthy for their control of the grade factor as a source of error. Studies in which this bias was uncontrolled have been reported by Crabbs (8), Barr *et al.* (3), Betts (4), and Taylor (25). In these latter studies, growth was measured upon a scattered sampling of classes from a range of grades.

It is reasonable to expect that part of a given teacher's effectiveness is attributable to what may be called a "school effect." Sources of this influence upon teacher performance, and consequently upon pupil growth, probably include the nature of the teacher's supervision, the adequacy of the physical plant, class-size policy, and the availability of materials and educational aids. The investigator, in attempting to identify differences among teachers, will need to incorporate in his design a method of identifying variance among schools as well as among the teachers within schools.

Studies included in this survey revealed lack of awareness of this problem. Those concerned only with teachers of one-room rural schools are perhaps somewhat less subject to this criticism than those in which several teachers from each of a number of heterogeneous schools are included. However, even in the case of rural-school studies, the data should be examined for specific biases produced by differing community conditions.

Almost every teacher believes that his pupils as a group possess certain tendencies or characteristics which either accelerate or retard growth. This kinds-of-pupils variable deserves the careful attention of the research worker. Unless pupil characteristics are controlled, the results of investigations have no general applicability.

It is common practice among school administrators to group pupils into classes on the basis of intelligence, social maturity, chronological age, or problem behavior. Although some of these

factors, like intelligence, can best be controlled in the analysis of data, the control of others must be provided in the experimental design. One scheme for experimentally controlling extraneous variables that enter into pupil growth is the random assignment of pupils to the classes of a given grade in a given school. Although random assignment may require much persuasion in the field, it must be conceded that effects due to pupil characteristics cannot be controlled without some randomization.

None of the studies reviewed in the present paper have indicated that pupil effects were controlled nor even that the necessity for control was recognized. The attempts of certain investigators to control their designs specifically for a few variables like socio-economic status, class size, and sex distribution is no substitute for randomization. Randomization is essential in order to spread the effects of *unknown* as well as *known* etiological factors, and thus to ensure the validity of whatever tests of significance are applied to the data.

THE last group of fundamental questions concerning pupil-growth research and its evaluation involves the choice of an analytic model appropriate to the data. The least adequate of the three techniques commonly used for the development of teacher-competence criteria based on pupil growth is raw gain. This technique assumes equality of units at every point in the score distribution and provides no means for controlling important sources of variation.

The major objection to the achievement-quotient technique rests on metrical grounds. The surrender of the information contained in pupils' raw scores for an age-equivalent ratio seems to be a wholly unnecessary sacrifice, since ratio scores can never yield more information than is contained in the raw scores from which they were derived. Johnson and Jackson (14) point out that adding, subtracting, multiplying, or dividing ratios such as $R=X/Y$, where Y varies from individual to individual, compounds errors of measurement. These arithmetic manipulations give rise to unjustifiable statistical inferences. Weaknesses of the achievement quotient are also discussed by Flanagan (10) and Cureton (9).

The technique applied in the majority of the recent studies of pupil growth is the multiple-regression method. As used in

the Wisconsin studies to arrive at index scores for teachers, the method is defective because the average gain score, the *predicted* variable, is not derived independently of the average initial score which is used as a *predictor* in the multiple-regression equation. For any given final score, a functional negative relationship must exist between a corresponding gain score and an initial score. In the Wisconsin studies the negative relationships found between pupil gain and pretest scores were attributed by the authors to "ceiling effects" of the tests employed. A more likely explanation is that the negative relationships are statistical artifacts based on the lack of independence of the two measures. This violation of statistical independence among the variates masks real differences that may exist among the means of classes. This defect in criterion development is characteristic of the studies reported by Rolfe (22), Rostker (23), Lins (17), Jones (15), Gotham (11), Von Eschen (27), and La Duke (16), despite their minor variations in the method of deriving residual pupil-gain scores.

Although the analysis of covariance has been available for thirty years, we have not found a single instance in which this technique was employed to secure a pupil-growth criterion. When its assumptions are fulfilled, the analysis of covariance is an eminently more appropriate model for establishing a pupil-growth criterion than either the multiple-regression or the achievement-quotient methods. The covariance technique enables the investigator to analyze the differences between the means of the final scores of a pupil measure for any number of classes and to adjust those means for the concomitant variation existing in the initial test score and for one or more other independent variables, such as mental ability. In this way covariance obviates the problem of lack of independence among the variables encountered in the use of residual-gain criteria obtained by multiple-regression methods. Moreover, covariance analysis is superior to multiple-regression analysis, because it takes into account all of the information contained in each individual score, whereas the latter method characteristically deals with class means. Adjustments in criterion means for independent variates are based on within-class regressions, rather than on between-class regressions as in multiple-regression solutions. If the relationship between dependent and independent variables is closer for class means than for individual pupils, then multiple regression "adjusts away" part of

the variation among classes that the technique is designed to estimate and therefore conceals part of the differences among teachers.

ALMOST all investigators have recognized the necessity for controlling the differences in intellectual ability of the pupils of those teachers to be compared. We do not expect dull children to make as much progress toward educational goals as bright children. Obviously, teachers of dull groups ought not to be handicapped, nor teachers of bright groups given an advantage in the comparison process.

The importance of controlling mental ability has been recognized by those research workers who used multiple-regression techniques, but there seems to be no particular point in partialing out the effects both of mental age *and* intelligence quotient, as Lins (17) has done. Still other investigators have proposed that the control of learning potential be accomplished by means of a complex matching technique. Johnson establishes the case against the use of a matching technique:

> Although the process of pairing serves to increase the sensitivity of the experimental observations, a considerable practical difficulty is encountered in obtaining representation throughout the range of variation. There is difficulty in securing a basic character that correlates highly enough with the criterion to ensure closely homogeneous groups. There is also considerable loss in information in that couples cannot be found for all individuals particularly if the standards of pairing are rigorous. In any case, selection is always a hazardous process since the nature of the sample becomes more uncertain as more or less arbitrary selection of individuals to be paired takes place. A further difficulty arises from the fact that the difference in the mean effect of the contrasted factors may depend upon the character or characters used in matching.[1]

If an investigator uses covariance technique and first partials out the effects of initial test on final test, additional control of a mental-ability score for each pupil may yield very slight further reductions in the several variance estimates. The factor governing whether elimination of the mental-ability effects results in a

[1] Barr, Arvil S.; Davis, Robert A.; and Johnson, Palmer O. *Educational Research and Appraisal*. Chicago: J. B. Lippincott Company, 1953, p. 233. (Lippincott Series in Education).

larger or smaller adjustment to the criterion depends upon the correlation of mental ability and initial test score. The closer this relationship is, the smaller the mental adjustment to the final test score will be. The covariance technique undoubtedly provides a way of treating mental ability which is more adaptable to the practical demands of educational situations than either the achievement-quotient or the multiple-regression technique.

Analysis of variance and covariance techniques provides an independent estimate of error in the form of the variation within the classes of pupils. The frequently used multiple-regression technique, however, offers no such built-in self evaluation since the variates are class means.

Having developed a pupil-growth criterion for assessing teaching effectiveness, the research worker should be expected to determine and report its power to discriminate among teachers. This power can be represented by an appropriate measure of the reliability of the growth criterion. Reliability is the chief quantitative aspect of any criterion, since the square root of the reliability coefficient represents the maximum correlation that can be found between even a perfectly reliable predictor variable and the criterion.

Failure to provide an estimate of the reliability of the growth index scores in discriminating among teachers is an almost universal shortcoming of the published research. There are two exceptions. Taylor (25) adapted the Spearman-Brown technique to his pupil-progress measures in reading and arithmetic by dividing the classes of pupils into random halves and correlating the gain scores of the halves. He found a corrected reliability coefficient of .415 in reading progress and .864 in arithmetic progress. Betts (4) used a similar method for his Stanford Achievement scores, estimating the reliability of the mean gain scores for 54 classes to be .895. The other researchers whose work we have reviewed either did not realize the significance of this question or expected their readers to accept their criteria on faith.

The multidimensionality of teaching competence makes the problem of assessing this area much more complex than originally believed by some investigators. Before a solution can be attempted, a specific hierarchy of educational goals and also adequate tests or other measuring devices for estimating pupil achievement must be devised.

Acceptance and Performance among Gifted Elementary-School Children

By META F. WILLIAMS

IT IS generally conceded that wasteful numbers of intellectually gifted children in public elementary schools do not perform to capacity in academic and social areas. Over the years, corrective measures such as acceleration, enrichment, and segregation have been tried and found wanting. Nevertheless, evidence has been accumulating to the effect that curricular changes can render the gifted child more accepting of, and more acceptable to, his classmates. This article summarizes the results of the writer's research to determine whether acceptance, in turn, can significantly influence performance.

In the initial step, three elementary schools of suburban Connecticut furnished a sample of 117 children whose intelligence quotient was 130 or more as derived from the California Tests of Mental Maturity. These children, deemed "gifted" for purposes of the experiment, were then rated for academic performance by Stanford Achievement Tests, Metropolitan Achievement Tests, and Metropolitan Readiness Tests (administered to kindergarten groups) and for social performance by the California Test of Personality.

To establish extent of acceptance, sociometric data were gathered by application of the Classroom Social Distance Scale.[1] This instrument allows each child to register on a five-point scale his reaction to every other child in the classroom. By assigning numerical values to the scale items, one can arrive at two scores: (1) a "self social distance" score indicating degree of acceptance extended to the group by an individual, and (2) a "group social distance" score denoting degree of acceptance extended to the individual by the group.

The scale was chosen, in part, because it assays both types of acceptance at issue, and also because it relieved the writer of the task of defining or analyzing "acceptance" except in terms of test results; thus, on the strength of the investigations connected with creation of the scale, the writer could logically construe

[1] Reprinted in Ruth Cunningham *et al.*, *Understanding Group Behavior of Boys and Girls* (New York: Bureau of Publications, Teachers College, Columbia University, 1951), pp. 401–406.

"friendliness" as a probable criterion of acceptance and, negatively, presume relative "rejection" to be a reasonable indication of relative lack of acceptance. The basis for these assumptions can be better comprehended upon examination of the scale items:

1. Would like to have him as one of my best friends.
2. Would like to have him in my group but not as a close friend.
3. Would like to be with him once in a while but not often or for long at a time.
4. Don't mind his being in our room but I don't want to have anything to do with him.
5. Wish he weren't in our room.[2]

With certain adjustments devised by the writer to accommodate children in kindergarten and the earlier grades, the scale was administered to 888 children, a total comprising the 117 gifted children and a random sample of their classmates.

After the sociometric results were completed, intensive statistical study was undertaken. Academic- and social-performance scores were analyzed by correlation ratio (Eta), by chi square, and by numerical comparison. Eta revealed a possible relation between acceptance extended *by the group* to the individual and academic as well as social performance. Also apparent was a relationship between acceptance extended *to the group* by the individual and the two types of performance.

Thus, academic performance yielded an Eta of .22 (significant at less than the .05 confidence level) in relation to acceptance extended by the group and an Eta of .58 (significant at greater than the .01 confidence level) in relation to acceptance extended to the group. Social performance correlated with acceptance by the group gave an Eta of .37 (significant at just less than the .05 confidence level) and with acceptance extended to the group an Eta of .30 (significant at less than the .05 confidence level).

The most pronounced Eta correlation, then, was between academic performance and acceptance extended to the group. The corrected contingency coefficient (Cc) brought out by chi square analysis likewise showed marked relation in this category, but the relation between social performance and acceptance extended to the group was still more marked. In general, Cc indicated that acceptance extended to the group was more closely related to

[2] *Ibid.*, p. 406.

varieties of performance than was acceptance extended by the group. Relation between total performance and total acceptance—the combined data of acceptance extended by the group and to the group—was found the most emphatic of all.

Table I shows the relation between two types of performance and two types of acceptance as indicated in statistical treatment

TABLE I

STATISTICAL RELATION BETWEEN PERFORMANCE AND ACCEPTANCE

Kinds of Relationship	Chi-Square	P	Corrected Contingency (Cc)
(1)	(2)	(3)	(4)
Academic performance and acceptance extended by the group	5.78	$.07 < P < .08$	.31
Academic performance and acceptance extended to the group	16.09	$P < .01$	.50
Social performance and acceptance extended by the group	9.07	$P < .01$	.41
Social performance and acceptance extended to the group	20.00	$P < .01$	.57
Total performance and acceptance extended by the group	14.76	$.01 < P < .02$	.43
Total performance and acceptance extended to the group	22.98	$P < .01$	.51
Total performance and total acceptance	41.39	$P < .01$	.62

by chi square. Corrected contingency coefficient (Cc) may be interpreted as signifying relative correlation between variables.[8] Total performance (academic plus social) and total acceptance (extended by the group and to the group) show the most pronounced relationship.

[8] *See* Charles C. Peters and Walter R. Van Voorhis, *Statistical Procedures and Their Mathematical Bases* (New York: McGraw-Hill Book Company, Inc., 1947); and J. P. Guilford, *Fundamental Statistics in Psychology and Education* (New York: McGraw-Hill Book Company, Inc., 1950).

Numerical analysis of the data revealed that more than four out of five children high in total acceptance were achieving within or beyond expectancy, whereas more than three out of five children low in acceptance were achieving below expectancy.

As a test of the findings, case studies were made of twelve children of the gifted sample, six of them highly accepted and six of them highly rejected. Techniques utilized, in addition to application of the scales previously mentioned, included the

TABLE II

COINCIDENCE OF EMOTIONAL NEED FULFILLMENT AND ACCEPTANCE EXTENDED BY THE GROUP

EMOTIONAL NEEDS	NUMBER OF HIGH ACCEPTEES IN WHOM NEED IS		NUMBER OF LOW ACCEPTEES IN WHOM NEED IS	
	Met	Unmet	Met	Unmet
(1)	(2)	(3)	(4)	(5)
Belonging	4	2	1	3
Achievement	4	2	2	4
Economic security	2	4	5	1
Freedom from fear	5	1	1	5
Love and affection	5	1	....	6
Freedom from excessive guilt	2	4	1	5
Self-respect	2	4	1	5
Self-understanding	6	....	2	4

sociometric interview, the anecdotal record of classroom observations, the school-record digest, open-end questions, and the Van Pit Series-Wishes.

The Van Pit Series-Wishes data, summarized in Table II, indicated need fulfillment among the intensively studied six gifted children highly accepted by classmates and the six gifted children rejected by classmates. The low acceptees showed greatest fulfillment in only one category, the need for *economic security*, which was met in five children of this group. The need for *love and affection* remained unfulfilled in all of the low acceptees; and in five cases each, the needs for *freedom from fear*, *freedom from excessive guilt*, and *self-respect* were not met. Among high acceptees, the need for *self-understanding* was fulfilled in every case. Of this group, five children had their needs for *freedom from fear* and *love and affection* met. The table indicates, then, a close relationship between acceptance and

satisfaction of needs. If schools make a greater effort to fulfill the needs of gifted children by strengthening acceptance, the writer concludes that, in all probability, the performance of such children will thereby be improved.

The case-study findings paralleled and supported those of the general investigation and, in addition, indicated the effects of the school setting and curriculum on degree of acceptance, and therefore, on quality of performance.

In summary, there appeared to be: (1) no appreciable differences in intelligence between high and low acceptees; (2) an appreciable difference in both test-measured and school-evaluated social performance, favoring the high acceptees; (3) evidence of greater acceptance extended to the group by the high acceptees; (4) no difference between high and low acceptees with regard to the characteristics valued in friends; (5) greater permissiveness and wider variety of channels for learning in the experiences of the high acceptees; (6) more consistent satisfaction with interpersonal relationships among the high, and more consistent dissatisfaction among the low acceptees; (7) considerable differences between high and low acceptees in the fulfillment of their emotional needs, as indicated by the Van Pit Series-Wishes.

Significantly, none of the individually studied gifted children, accepted by classmates or not, appeared concerned with himself as a "big brain." None sought special attention on the grounds of unusual intellectual capacity. None was anxious for more school work or for special teachers or for a rarified climate of any kind. The profuse data of the individual studies also indicated that all of the examined children, accepted by classmates or not, were willing, within emotional limitations, to attempt the work prescribed by the school.

Moreover, all were powerfully concerned with themselves as children, not as special children. They indicated the normal strong desire of children to use their bodies successfully, to build themselves physically into greater beauty and strength. They all wanted to love and be loved. And they wanted, above all, to consider themselves worthy members of a worthy peer group and to communicate this desire to their classmates. It seems reasonable to conclude, then, that those engaged in seeking better opportunities for our gifted children might with profit look to the child first, and only subsequently to his gifts.

[*Continued on page 224*]

EDITORIAL COMMENT

A Year of the Space Age

A YEAR has passed since the launching of Sputnik I shocked the public and the educational profession out of their complacent assumptions that in science, technology, and education America was, as a matter of course, ahead of the rest of the world, and that we had no reason to be seriously concerned about competition from the Soviet Union. As might have been expected, this complacency was followed by near-hysteria in certain quarters. We are now far enough away from the first shock for a relatively calm appraisal of events since the launching. It is appropriate now to ask, what progress has the year brought in making our education more nearly adequate to meet the needs of the age of satellites, missiles, explosive nationalism, and Communist imperialism?

It seems safe to say that, educationally, Sputnik I was a blessing in disguise. More has been accomplished than would have been possible without the stimulus that it provided. There has been substantial progress in meeting the need for schoolhouse construction. The shortage of classrooms is still severe, but it is being reduced. Teachers' salaries, although still shamefully low, have been improved. Much has been done by the National Science Foundation and other organizations and institutions to strengthen the preparation of secondary-school teachers, especially in the areas of mathematics and science. More than three hundred colleges and universities are now participating in a televised college course in Atomic-Age Physics for high-school science teachers, which began operation on October 6. For the first time a course carrying college credit is being given nationally on television. It promises to make an important contribution to the improvement of science teaching.

The most conspicuous item in the year's advance was the enactment of the National Defense Education Act of 1958 (the Hill-Elliott Act). It is an omnibus measure providing, among other things, for loans to needy, able students; increased support for the teaching of science, mathematics, and foreign languages; foreign-language institutes to improve the competence of language teachers; fellowships for students in new or expanded

graduate programs; funds to assist the states in establishing and maintaining guidance and counseling programs; and support of research in the development and use of audio-visual aids.[1] This act promises to have far-reaching influence in the fields with which it deals. Almost certainly it could not have been passed without the shock of Sputnik.

THERE has been similar progress in less tangible matters. Many observers report a more serious attitude toward intellectual concerns on the part of students and their parents. The word *egghead* is much less popular than it was; the good student and the intellectual are regarded with growing respect. There is a more critical attitude toward the high-school curriculum and a definite decline in the tendency to regard any subject as valuable as any other if it is equally well taught for the same length of time. Providing superior students with educational experiences that will challenge them and contribute most effectively to their intellectual development is being increasingly emphasized.

These and many other developments that might be mentioned add up to substantial progress in the first year of the space age. They do not, however, justify complacency. Much remains to be done along many lines; we are only well started on the job. The most basic need of all is that of developing and maintaining a sense of urgency on the part of the public, the profession, and the students.

R. H. E.

* * *

The findings of a recent study of high-school students in Ohio enrolled in mathematics and science courses have been summarized in *Ohio Schools*, October, 1958. Significantly, nearly 10 per cent more of today's students are taking such courses than in 1947. Of the total enrollment, 52 per cent are taking a mathematics course and 58.7 per cent, a science course. Enrollment in "watered down" science courses, such as general science, is "negligible," only 1.07 per cent. Seniors are electing the "tough" college preparatory courses like solid geometry, trigonometry, and senior mathematics (18.7 per cent); and physics, anatomy, zoology, and senior science (27.25 per cent).

[1] We hope to discuss this act more fully in a future issue of the EDUCATIONAL RESEARCH BULLETIN.

READINGS

Books to Read

WARNER, RUBY H. *The Child and His Elementary School World.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1957. x+406 pp.

In the opening chapter, "The Self and the Group," the author turns to the pledge of the Mid-century White House Conference on Children and Youth for the five vital needs of all children—"affection, a feeling of belonging, cooperative living, creating, and a sense of achievement" (page 3). These she carefully amplifies in terms of meeting the psychological, biological, sociological, and educational needs as they intermix and shape the development of self and groups. The discussion about meeting educational needs proceeds from a philosophical base and examines in turn the current thinking about school organization, curriculum, achievement, and evaluation, in terms of furthering self-realization and better human relations for children.

Throughout the next eight chapters of the book the author faithfully traces the many "musts" of educational content which become meaningful learning content in programs designed to meet the vital needs of children. This she does through summarizing the potential contribution of an area, such as language arts, to the desirable growth of individuals and groups. Then, by illustrating and analyzing particular situations in terms of procedures and outcomes, the author makes clear the child's rôle as an active agent rather than a passive recipient in his school experiences. In the chapters about "Creative Writing and the Self," "Creative Art and the Self," and "Music and the Self," the emphasis is upon the creative aspects of these areas in the development of personality. Science is presented as "relationship with the environment" and "Arithmetic as Quantitative Thinking." "Health, Physical Education, and Safety" are dealt with in a special chapter. The chapter which makes clear the interrelationships of the various areas is the one on "Social Education for Group Living." The contributions of these areas to problem solving become apparent in the accounts of social studies units for kindergarten through Grade VI.

In the final chapter, "Home, School and Community," five of the many ties among these groups which call for co-operative effort in nurturing young citizens are outlined: the parent-teacher association, home-school visits, resources in the community, field trips, and reporting to parents. According to the author,

> The parent-teacher conference has three important features—planning, conferring, and helping the child through mutual guidance.
>
> . . . . . .
>
> From information offered by the parent, the teacher may be saved

valuable time in becoming oriented to the child's abilities. One parent helped a beginning teacher understand the attractive personality and limited capabilities of her daughter, a newcomer in the school (page 390).

The format of this book lends itself to easy reference. Illustrations of classroom procedures at various grade levels are followed by the author's analysis of the situation. Brief italicized summaries and selected bibliographies are included after each chapter. These factors, along with the author's talent for interlacing theory and practice, add up to an important contribution to elementary education.

MARY JANE LOOMIS

* * *

The Student Personnel Program

[*Continued from page 204*]

Evaluation may precede and planning for improvement may follow, but research is the process which places future developments on a respectable basis of fact rather than fancy. At the present time two modest projects are under way: (1) an experiment to determine criteria for reinstating dismissed students and (2) a follow-up study of a group of Freshmen who enrolled last fall with scholarships in teacher education.

If one were to evaluate the present student personnel program, he would surely pinpoint numerous questions to be investigated and several procedures to be restructured. Perhaps the present effectiveness of the program is due not so much to what is done, as to how it is done and by whom. [Vol. XXXVII, No. 8]

Acceptance and Performance

[*Continued from page 220*]

If further investigation can demonstrate that the performance of gifted children is generally affected by considerations of group acceptance, as seemed true of the children in the study reported here, then the school must re-examine those practices which are inimical to acceptance and change them. Such change, aimed at cultivating the child rather than his particular gifts, might well prove less expensive and more effective than the practices now relied upon to help intellectually endowed children attain performance commensurate with their special capacities.

[Vol. XXXVII, No. 8]